



ILLUSTRATION BY ANDY POTTS

Will an astronaut who falls into a black hole be crushed or burned to a crisp?

FIRE IN THE HOLE!

BY ZEEYA MERALI

I

In March 2012, Joseph Polchinski began to contemplate suicide — at least in mathematical form. A string theorist at the Kavli Institute for Theoretical Physics in Santa Barbara, California, Polchinski was pondering what would happen to an astronaut who dived into a black hole. Obviously, he would die. But how?

According to the then-accepted account, he wouldn't feel anything special at first, even when his fall took him through the black hole's event horizon: the invisible boundary beyond which nothing can escape. But eventually — after hours, days or even weeks if the black hole was big enough — he would begin to notice that gravity was tugging at his feet more strongly than at his head. As his plunge carried him inexorably downwards, the difference in forces would quickly increase and rip him apart, before finally crushing his remnants into the black hole's infinitely dense core.

But Polchinski's calculations, carried out with two of his students — Ahmed Almheiri and James Sully — and fellow string theorist Donald Marolf at the University of California, Santa Barbara (UCSB), were telling a different story¹. In their account, quantum effects would turn the event horizon into a seething maelstrom of particles. Anyone who fell into it would hit a wall of fire and be burned to a crisp in an instant.

The team's verdict, published in July 2012, shocked the physics community. Such firewalls would violate a foundational tenet of physics that was first articulated almost a century ago by Albert Einstein, who used it as the basis of general relativity, his theory of gravity. Known as the equivalence principle, it states in part that an observer falling in a gravitational field — even the powerful one inside a black hole — will see exactly the same phenomena as an observer floating in empty space. Without this principle, Einstein's framework crumbles.

Well aware of the implications of their claim, Polchinski and his co-authors offered an alternative plot ending in which a firewall does not form. But this solution came with a huge price. Physicists would have to sacrifice the other great pillar of their science: quantum mechanics, the theory governing the interactions between subatomic particles.

The result has been a flurry of research papers about firewalls, all struggling to resolve the impasse, none succeeding to everyone's satisfaction. Steve Giddings, a quantum physicist at the UCSB, describes the situation as "a crisis in the foundations of physics that may need a revolution to resolve".

With that thought in mind, black-hole experts came together last month at CERN, Europe's particle-physics laboratory near Geneva, Switzerland, to grapple with the issue face to face. They hoped to reveal the path towards a unified theory of 'quantum gravity' that brings all the fundamental forces of nature under one umbrella — a prize that has eluded physicists for decades.

The firewall idea "shakes the foundations of what most of us believed about black holes", said Raphael Bousso, a string theorist at the University of California, Berkeley, as he opened his talk at the meeting. "It essentially pits quantum mechanics against general relativity, without giving us any clues as to which direction to go next."

FIERY ORIGINS

The roots of the firewall crisis go back to 1974, when physicist Stephen Hawking at the University of Cambridge, UK, showed that quantum effects cause black holes to run a temperature². Left in isolation, the holes will slowly spew out thermal radiation — photons and other particles — and gradually lose mass until they evaporate away entirely (see 'The information paradox').

These particles aren't the firewall, however; the subtleties of relativity guarantee that an astronaut falling through the event horizon will not notice this radiation. But Hawking's result was still startling — not least because the equations of general relativity say that black holes can only swallow mass and grow, not evaporate.

Hawking's argument basically comes down to the observation that in the quantum realm, 'empty' space isn't empty. Down at this sub-sub-microscopic level, it is in constant turmoil, with pairs of particles and their corresponding antiparticles continually popping into existence before rapidly recombining and vanishing. Only in very delicate laboratory experiments does this submicroscopic frenzy have any observable consequences. But when a particle–antiparticle pair appears just outside a black hole's event horizon,

Hawking realized, one member could fall in before the two recombined, leaving the surviving partner to fly outwards as radiation. The doomed particle would balance the positive energy of the outgoing particle by carrying negative energy inwards — something allowed by quantum rules. That negative energy would then get subtracted from the black hole's mass, causing the hole to shrink.

Hawking's original analysis has since been refined and extended by many researchers, and his conclusion is now accepted almost universally. But

"IT IS A CRISIS IN THE FOUNDATIONS OF PHYSICS."

with it came
the disturbing realization that
black-hole radiation leads to a paradox that chal-
lenges quantum theory.

Quantum mechanics says that information cannot be destroyed. In principle, it should be possible to recover everything there is to know about the objects that fell in a black hole by measuring the quantum state of the radiation coming out. But Hawking showed that it was not that simple: the radiation coming out is random. Toss in a kilogram of rock or a kilogram of computer chips and the result will be the same. Watch the black hole even until it dies, and there would still be no way to tell how it was formed or what fell in it.

This problem, dubbed the black-hole information paradox, divided physicists into two camps. Some, like Hawking, argued that the information truly vanishes when the black hole dies. If that contradicted quantum laws, then better laws needed to be found. Others, like John Preskill, a quantum physicist at the California Institute of Technology in Pasadena, stuck by quantum mechanics. "For a time, I did seriously try to build an alternative theory that included information loss," he says. "But I couldn't find one that made any sense — nobody could." The stalemate continued for the next two decades, finding its most famous expression in 1997, when Preskill publicly bet Hawking that information was not being lost, with the winner to receive an encyclopaedia of his choice.

But that same year, the deadlock was broken by a discovery made by Juan Maldacena, a physicist then at Harvard University in Cambridge. Maldacena's insight built on an earlier proposal that any three-dimensional (3D) region of our Universe can be described by information encoded on its two-dimensional (2D) boundary^{3–5}, in much the same way that laser light can encode a 3D scene on a 2D hologram. "We used the word 'hologram' as a metaphor," says Leonard Susskind, a string theorist at Stanford University in California, and one of those who came up with the proposal⁴. "But after doing more mathematics, it seemed to make literal sense that the Universe is a projection of information on the boundary."

What Maldacena came up with was a concrete mathematical

formulation⁶ of the hologram idea that made use of ideas from superstring theory, which posits that elementary particles are composed of tiny vibrating loops of energy. His model envisages a 3D universe containing strings and black holes that are governed only by gravity, bounded by a 2D surface on which elementary particles and fields obey ordinary quantum laws without gravity. Hypothetical residents of the 3D space would never see this boundary because it is infinitely far away. But that wouldn't matter: anything happening in the 3D universe could be described equally well by equations in the 2D universe, and vice versa. "I found that there's a mathematical dictionary that allows you to go back and forth between the languages of these two worlds," Maldacena explains.

This meant that even 3D black-hole evaporation could be described in the 2D world, where there is no gravity, where quantum laws reign supreme and where information can never be lost. And if information is preserved there, then it must also be preserved in the 3D world. Somehow, information must be escaping from the black holes.

ONE FOR ALL

A few years later, Marolf showed that every model of quantum gravity will obey the same rules, whether or not it is built from string theory⁷. "It was a combination of Maldacena and Marolf's work that turned me around," explains a long-term proponent of information loss, Ted Jacobson, a quantum physicist at the University of Maryland in College Park. In 2004, Hawking publicly admitted that he had been wrong, and gave

For a particle to be emitted at all, it has to be entangled with the twin that is sacrificed to the black hole. And if Susskind and others were right, it also had to be entangled with all the Hawking radiation emitted before it. Yet a rigorous result of quantum mechanics dubbed 'the monogamy of entanglement' says that one quantum system cannot be fully entangled with two independent systems at once.

To escape this paradox, Polchinski and his co-workers realized, one of the entanglement relationships had to be severed. Reluctant to abandon the one required to encode information in the Hawking radiation, they decided to snip the link binding an escaping Hawking particle to its infalling twin. But there was a cost. "It's a violent process, like breaking the bonds of a molecule, and it releases energy," says Polchinski. The energy generated by severing lots of twins would be enormous. "The event horizon would literally be a ring of fire that burns anyone falling through," he says. And that, in turn, violates the equivalence principle and its assertion that free-fall should feel the same as floating in empty space — impossible when the former ends in incineration. So they posted a paper on the pre-print server, arXiv, presenting physicists with a stark choice: either accept that firewalls exist and that general relativity breaks down, or accept that information is lost in black holes and quantum mechanics is wrong¹. "For us, firewalls seem like the least crazy option, given that choice," says Marolf.

The paper rocked the physics community. "It was outrageous to claim that giving up Einstein's equivalence principle is the best option," says Jacobson. Bousso agrees, adding: "A firewall simply can't appear in empty space, any more than a brick wall can suddenly appear in an empty field and smack you in the face." If Einstein's theory doesn't apply at the event horizon, cosmologists would have to question whether it fully applies anywhere.

Polchinski admits that he thought they could have made a silly mistake. So he turned to Susskind, one of the fathers of holography, to find it. "My first reaction was that they were wrong," says Susskind.

He posted a paper stating as much⁸, before quickly retracting it, after further thought. "My second reaction was that they were right, my third was that they were wrong again, my fourth was that they were right," he laughs. "It's earned me the nickname, 'the yo-yo,' but my reaction is pretty much the same as most physicists."

Since then, more than 40 papers have been posted on the topic in arXiv, but as yet, nobody has found a flaw in the team's logic. "It's a really beautiful argument proving that there's something inconsistent in our thinking about black holes," says Don Page, a collaborator of Hawking's during the 1970s who is now at the University of Alberta in Edmonton, Canada. A number of inventive solutions have been offered, however.

REAL-WORLD IMPLICATIONS

One of the most promising resolutions, according to Susskind, has come from Daniel Harlow, a quantum physicist at Princeton University in New Jersey, and Patrick Hayden, a computer scientist at McGill University in Montreal, Canada. They considered whether an astronaut could ever detect the paradox with a real-world measurement. To do so, he or she would first have to decode a significant portion of the outgoing Hawking radiation, then dive into the black hole to examine the infalling particles. The pair's calculations show that the radiation is so tough to decode that the black hole would evaporate before the astronaut was ready to jump in⁹. "There's

"THE EVENT HORIZON IS LITERALLY A RING OF FIRE."

Preskill a baseball encyclopaedia to make good on their bet.

Such was the strength of Maldacena's discovery that most physicists believed that the paradox had been settled — even though nobody had yet explained how Hawking radiation smuggles information out of the black hole. "I guess we just all assumed there would be a straightforward answer," says Polchinski.

There wasn't. When Polchinski and his team set themselves the task of clearing up that loose end in early 2012, they soon stumbled on yet another paradox — the one that eventually led them to the fatal firewall.

Hawking had shown that the quantum state of any one particle escaping from the black hole is random, so the particle cannot be carrying any useful information. But in the mid-1990s, Susskind and others realized that information could be encoded in the quantum state of the radiation as a whole if the particles could somehow have their states 'entangled' — intertwined in such a way that measurements carried out on one will immediately influence its partner, no matter how far apart they are.

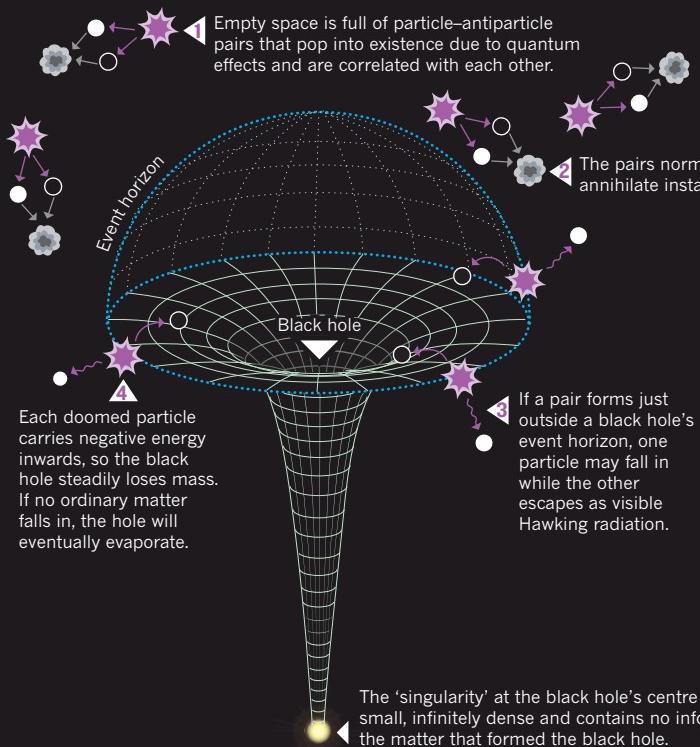
But how could that be, wondered the Polchinski's team?

THE INFORMATION PARADOX

Matter falling into a black hole is crushed to an infinitely dense point at the centre. Two scenarios attempt to explain what happens to the information that matter holds.

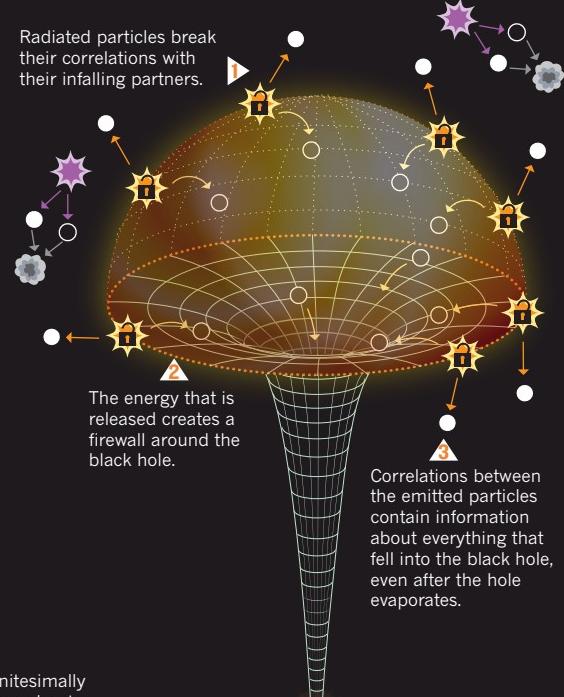
DISAPPEARANCE

When the hole evaporates, all the information disappears with it.



FIREWALL

The information is carried out by quantum correlations among all particles radiated from the black hole.



no fundamental law preventing someone from measuring the paradox," says Harlow. "But in practice, it's impossible."

Giddings, however, argues that the firewall paradox requires a radical solution. He has calculated that if the entanglement between the outgoing Hawking radiation and its infalling twin is not broken until the escaping particle has travelled a short distance away from the event horizon, then the energy released would be much less ferocious, and no firewall would be generated¹⁰. This protects the equivalence principle, but requires some quantum laws to be modified. At the CERN meeting, participants were tantalized by the possibility that Giddings' model could be tested: it predicts that when two black holes merge, they may produce distinctive ripples in space-time that can be detected by gravitational-wave observatories on Earth.

There is another option that would save the equivalence principle, but it is so controversial that few dare to champion it: maybe Hawking was right all those years ago and information is lost in black holes. Ironically, it is Preskill, the man who bet against Hawking's claim, who raised this alternative, at a workshop on firewalls at Stanford at the end of last year. "It's surprising that people are not seriously thinking about this possibility because it doesn't seem any crazier than firewalls," he says — although he adds that his instinct is still that information survives.

The reluctance to revisit Hawking's old argument is a sign of the immense respect that physicists have for Maldacena's dictionary relating gravity to quantum theory, which seemingly proved that information cannot be lost. "This is the deepest ever insight into gravity because it links it to quantum fields," says Polchinski, who compares Maldacena's

result — which has now accumulated close to 9,000 citations — to the nineteenth-century discovery that a single theory connects light, electricity and magnetism. "If the firewall argument had been made in the early 1990s, I think it would have been a powerful argument for information loss," says Bousso. "But now nobody wants to entertain the possibility that Maldacena is wrong."

Maldacena is flattered that most physicists would back him in a straight-out fight against Einstein, although he believes it won't come to that. "To completely understand the firewall paradox, we may need to flesh out that dictionary," he says, "but we won't need to throw it out."

The only consensus so far is that this problem will not go away any time soon. During his talk, Polchinski fielded all proposed strategies for mitigating the firewall, carefully highlighting what he sees as their weaknesses. "I'm sorry that no one has gotten rid of the firewall," he concludes. "But please keep trying." ■

Zeeya Merali is a freelance writer based in London.

1. Almheiri, A., Marolf, D., Polchinski, J. & Sully, J. Preprint at <http://arxiv.org/abs/1207.3123> (2012).
2. Hawking, S. W. *Nature* **248**, 30–31 (1974).
3. Bekenstein, J. D. *Phys. Rev. D* **7**, 2333–2346 (1973).
4. Susskind, L. *J. Math. Phys.* **36**, 6377 (1995).
5. Stephens, C. R., 't Hooft, G. & Whiting, B. F. *Class. Quant. Grav.* **11**, 621–647 (1994).
6. Maldacena, J. M. *Adv. Theor. Math. Phys.* **2**, 231–252 (1998).
7. Marolf, D. *Phys. Rev. D* **79**, 044010 (2009).
8. Susskind, L. Preprint at <http://arxiv.org/abs/1207.4090> (2012).
9. Harlow, D. & Hayden, P. Preprint at <http://arxiv.org/abs/1301.4504> (2013).
10. Giddings, S. B. Preprint at <http://arxiv.org/abs/1302.2613> (2013).